AD_____

Award Number: DAMD17-01-1-0179

TITLE: Non-Invasive Screening for Breast Cancer Biomarkers in Nipple Aspirates by Mass Spectrometry

PRINCIPAL INVESTIGATOR: Helena R. Chang, M.D., Ph.D.

CONTRACTING ORGANIZATION: The University of California,
Los Angeles
Los Angeles, California 90024-1406

REPORT DATE: July 2002

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20030313 184

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2002	Annual (18 Jun 01 - 17 Jun 02)

4. TITLE AND SUBTITLE
Non-Invasive Screening for Breast Cancer Biomarkers in Nipple Aspirates by Mass Spectrometry

5. FUNDING NUMBERS
DAMD17-01-1-0179

6. AUTHOR(S)
Helena R. Chang, M.D., Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The University of California, Los Angeles
Los Angeles, California 90024-1406

E-Mail: hchang@mednet.ucla.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

14. SUBJECT TERMS
Breast Cancer

15. NUMBER OF PAGES
6

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Specific Aims

The three challenges that we faced before are:

1. To consistently differentiate peak differences in NA detected by Surface-Enhance Laser Desorption/Ionization (SELDI) between the two groups of women: breast cancer vs. non-cancer.
2. To identify and characterize the peak proteins that are differentially expressed in cancerous and non-cancerous NA.
3. To develop software program that can consistently identify the true peak from noise and quantify the intensity of each peak in a reproducible fashion.

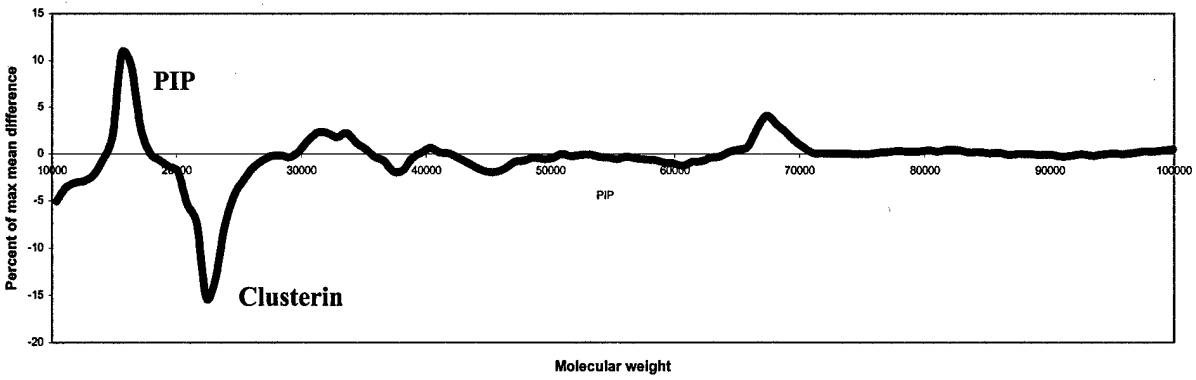
Results

1. Using SELDI mass spectrometry we are able to generate from each breast cancer specimen a semi-quantitative profile of proteins. Currently, we have analyzed 88 nipple aspirates derived from cancerous (47) and noncancerous (41) breasts. Eleven protein peaks are shown to be statistically different between the two groups: cancer vs. non-cancer. The results are summarized below:

Table: Molecular weights of protein peaks that differentiate cancerous from non-cancerous nipple aspirates (NA)

Molecular Weight (dalton)	Mean std Intensity diff	T	P
10370	-5.057	2.24	0.0254
11118	-3.646	1.61	0.1070
15618	10.76	-4.76	<.0001
16368	9.422	-4.16	<.0001
20868	-5.392	2.38	0.0172
21618	-7.323	3.24	0.0012
22368	-15.19	6.71	<.0001
23118	-13.31	5.88	<.0001
23868	-7.743	3.42	0.0006
24618	-4.442	1.96	0.0503
67368	4.085	-1.81	0.0710

2. The protein ID of the two prominent protein peaks by protein separation, cleavage and amino acid sequencing by mass spectrometry suggested that the PIP (prolactin inducing protein) is abundant in nipple aspirates of cancerous breasts and clustering is abundant in the nipple aspirates of noncancerous breasts. We are confident our collaborators, Drs. Kym and Whitelegge will extend their expertise in helping us to design and refine the proteomic analysis.



3. Statistical Analysis – Software program development and comparison
The SBCC staff (Jeff Gornbein & Rita Engelhardt, with input from Dr. Elliot Landaw) has been exploring the problem of identifying peaks from SELDI.

The preliminary analyses we gave showed that the standardized protein profiles are reasonably well modeled by a Gaussian, even without scale transformations such as log transformations. We have the capability to generalize beyond the classical repeated measure assumptions as our model (SAS procedure MIXED) allows for variance heterogeneity and can allow for general patterned within and between subject covariate matrices.

Since we plan to limit the range of our protein profile comparisons to the 10,000 to 100,000 dalton range and since we can adjust size of the "bins" used to partition this range, we generally do not have a problem with missing data in any bin for any one profile. In fact, since intensity observations are made approximately every 0.2 dalton, we typically have over 500,000 observations per profile. The large number of observations per bin also helps explain why a Gaussian model is reasonable, as distributions tend toward the Gaussian when n is large, particularly when we average over all the observations in each bin for each profile. Moreover, our model uses maximum likelihood methods that allow use of all the observed data even if there was missing data for some bins in some samples, provided that the data is missing at random.

Due to current computer memory limitations, we then compress the data into bins that are 150 data points "wide". Since there are about 5 daltons per data point, this bin is about 750 daltons wide. Thus, the current resolution is only to about 750/2 = 375 daltons. This can easily be improved with a more powerful computer allowing for smaller bins. The current bin size on a Pentium II compute with 300 MHz takes 3 minutes to run.

The results presented in this report were generated through analysis using relative peak heights for calculation. Another approach we could utilize is comparing peak area as criteria for

analysis. We plan to calculate the relative peak area profiles of cancerous and non-cancerous NA when a more powerful computer becomes available.